Iterative solution of linear equations with unbounded operators

A.G. Ramm
Mathematics Department, Kansas State University,
Manhattan, KS 66506-2602, USA
ramm@math.ksu.edu
http://www.math.ksu.edu/~ramm

Abstract

A convergent iterative process is constructed for solving any solvable linear equation in a Hilbert space.

1 Introduction

A basic general result about solvable linear equations

$$Au = f, \tag{1}$$

where A is a linear bounded operator in a Hilbert space, is the following theorem.

Theorem 0. *Any solvable equation* (1) *with a bounded linear operator can be solved by a convergent iterative process.*

A proof of Theorem 0 can be found, e.g., in [2]. One of the steps in this proof is the following simple Lemma (see e.g. [2]).

Lemma 1. *If equation* (1) *is solvable and A is a bounded linear operator, then equation* (1) *is equivalent to*

$$A^*Au = A^*f. \tag{2}$$

Math subject classification: 44A05 47A50 65J10
key words: iterative methods, unbounded operators, linear equations.

The equivalence in Lemma 1 means that every solution to (1) solves (2) and vice versa.

The aim of this paper is to study equation (1) with a linear, closed, densely defined, unbounded, and not necessarily boundedly invertible operator. In other words, A may be not injective, i.e., null space $N := N(A)$ may be non-trivial, and its range $R(A)$ may be not closed. Although there are many papers and books on iterative methods, iterative methods for equations (1) with unbounded operators were not studied in such generality.

Our second aim is to study a variational regularization method for the solutions to equation (1). By y we denote throughout the unique solution to (1) of minimal norm, i.e., the solution $y \perp N$. This solution will be of main interest to us. If A is bounded, but not boundedly invertible, so that (1) is an ill-posed problem (see e.g. [2]), then a variational regularization method for obtaining a stable approximation of the solution y given noisy data f_δ, $\|f_\delta - f\| \le \delta$, consists of

a) minimizing the functional

$$F(u) = \|Au - f_\delta\|^2 + a\|u\|^2, \qquad a > 0, \tag{3}$$

where a is a constant called a regularization parameter, proving that (3) has a unique global minimizer $u_{a,\delta} = (A^*A + aI)^{-1}A^*f_\delta$,

and

b) proving that one can choose $a = a(\delta)$, so that $\lim_{\delta\to 0} a(\delta) = 0$ and

$$\lim_{\delta\to 0} \|u_\delta - y\| = 0, \qquad u_\delta := u_{a(\delta),\delta}. \tag{4}$$

Formula (4) shows that u_δ is a stable approximation of y. The rate of convergence of u_δ to y is not possible to specify without imposing additional assumptions on f.

If A is unbounded, then it was not proved that functional (3) has a unique global minimizer. Formula $u_{a,\delta} = (A^*A + aI)^{-1}A^*f_\delta$ is not well defined because f_δ may not belong to $D(A^*)$.

Throughout the paper $T = A^*A$ is a selfadjoint nonnegative operator (generated by the closed nonnegative quadratic form (Au, Au), $D(T) \subset D(A)$, $T_a := T + aI$, I is the identity operator, $Q = AA^* \ge 0$ is a selfadjoint operator, $D(Q) \subset D(A^*)$. Recall that A^* is well defined if A is densely defined, and A^* is densely defined if A is closed. (See [1]). By S_0 we denote the operator $T_a^{-1}A^*$ with domain $D(A^*)$, and by S we denote its closure.

Our results can be described as follows: the operator S_0 is closable, its closure is defined on all of H and is a bounded operator, $\|S\| \le \frac{1}{2\sqrt{a}}$. A similar result holds for $S_1 = T_{ia}^{-1}A^*$ with domain $D(A^*)$. Our result shows that the element $T_a^{-1}A^*f_\delta$ is well defined for any $f_\delta \in H$, and not only for $f_\delta \in D(A^*)$.

Consider the iterative process

$$u_{n+1} = Bu_n + T_a^{-1}A^*f, \quad u_1 = u_1, \quad a > 0, \tag{5}$$

where the initial approximation $u_1 \perp N$ and otherwise arbitrary, and $B := aT_a^{-1}$.

Theorem 1. *If A is a linear, closed, densely defined operator in H, $a > 0$, $B = aT_a^{-1}$, and $u_1 \perp N$, then*

$$\lim_{n\to\infty} \|u_n - y\| = 0, \tag{6}$$

where u_n is defined by (5).

Theorem 2. *If A is a linear, closed, densely defined operator in H, and $a > 0$ is a constant, then the operator $S_0 = T_a^{-1}A^*$ with domain $D(A^*)$ is closable and its closure S is a bounded operator defined on all of H, $\|T_a^{-1}A^*\| \leq \frac{1}{2\sqrt{a}}$. Similar results hold for the operator $T_{ia}^{-1}A^*$.*

Theorem 3. *For any $f_\delta \in H$ functional* (3) *has a unique global minimizer $u_{a,\delta} = T_a^{-1}A^*f_\delta = A^*Q_a^{-1}f_\delta$.*

In Section 2 proofs are given. In Section 3 we construct a stable approximation to y given noisy data f_δ and using an iterative process similar to (5). In Section 4 the case of selfadjoint, unbounded and possibly not boundedly invertible operator is briefly considered. In Section 5 the dynamical systems method (DSM) (developed in [2] pp.41-70) is justified for equation (1) with unbounded, linear, densely defined operator in a Hilbert space. The basic results of this paper are stated in Theorems 1 through 5.

2 Proofs

Proof of Theorem 2. To prove the closability of S_0, assume that $u_n \in D(A^*) = D(S_0)$, $u_n \to 0$, $S_0 u_n \to f$, and prove $f = 0$. We have

$$(f, h) = \lim_{n\to\infty}(S_0 u_n, h) = \lim_{n\to\infty}(A^*u_n, T_a^{-1}h) = \lim_{n\to\infty}(u_n, AT_a^{-1}h) = 0, \qquad \forall h \in H. \tag{7}$$

Here we have used the inclusion $R(T_a^{-1}) \subset D(A)$. This inclusion can be verified: if $g = T_a^{-1}h$, then $Tg + ag = h$, so $g \in D(T) \subset D(A)$, as claimed. From (7) it follows that $f = 0$ because $h \in H$ is arbitrary. Relation (7) shows that $D(S_0^*) = H$ and $S_0^* = AT_a^{-1}$. This operator is closed and densely defined. Indeed, by the polar decomposition, $A = UT^{\frac{1}{2}}$, where U is a partial isometry, $\|U\| \leq 1$. The operator $T^{\frac{1}{2}}T_a^{-1}$ is densely defined, it is a function of the selfadjoint operator T. We have

$$\|S_0^*\| \leq \|T^{\frac{1}{2}}T_a^{-1}\| = \|\int_0^\infty \frac{s^{\frac{1}{2}}}{s+a} dE_s\| = \sup_{s\geq 0} \frac{s^{\frac{1}{2}}}{s+a} = \frac{1}{2\sqrt{a}}, \tag{8}$$

where we have used the spectral theorem and E_s is the resolution of the identity of the selfadjoint operator T. Since $\|S\| = \|S_0^{**}\| = \|S_0^*\| \leq \frac{1}{2\sqrt{a}}$, Theorem 2 is proved except for the claim concerning the operator $T_a^{-1}A^*$. The proof of this claim is essentially the same as the above proof, the only (not important) difference is the replacement of the formula $(T_a^{-1})^* = T_a^{-1}$ by $(T_{+ia}^{-1})^* = T_{-ia}^{-1}$.

Theorem 2 is proved. $\square$

Proof of Theorem 1. If equation (1) is solvable then $f = Ay$, the operator $B = aT_a^{-1}$ is bounded, defined on all of H, and $\|B\| \leq 1$.

Consider the equation

$$u = Bu + T_a^{-1}A^*f. \tag{9}$$

This equation makes sense for any $f \in H$ by Theorem 2. The minimal-norm solution y to equation (1) solves (9) in the following sense.

If $f \in D(A^*)$, then $Ty = A^*f$, $ay + Ty = ay + A^*f$, $y = By + T_a^{-1}A^*f$, so y solves (9) if $f \in D(A^*)$. Since the set $f \in D(A^*) \bigcap R(A) = D(T)$ is dense in H and, consequently, in $D(A^*)$, and since, by Theorem 2, the operator $S = T_a^{-1}A^*$ is uniquely extendable to all of H (from a dense subset $D(A^*)$) by continuity, it follows that if y solves equation (9) for every $f \in D(A^*) \bigcap R(A)$, then this equation is solvable for any $f \in R(A)$. Indeed, suppose $u_n = Bu_n + Sf_n$, $f_n \in \underline{D(A)} \bigcap R(A)$, $u_n \perp N$, and $\lim_{n\to\infty} f_n = f \in R(A)$. The subspace $N^\perp := N(T)^\perp = \overline{R(T)}$ is invariant with respect to B. If $f \in R(A)$ and $f = Ay$, then equation (9) has a unique solution $u \in N^\perp$, and this solution is $u = (I - B)^{-1}T_a^{-1}Ty = y$.

Indeed,

$$(I - B)^{-1}T_a^{-1}Ty = (I - aT_a^{-1})^{-1}T_a^{-1}Ty = \int_0^\infty \frac{sdE_sy}{(1 - \frac{a}{a+s})(s + a)} = y$$

Denote $w_n := u_n - y$. Then (5) and equation (9) for y imply $w_{n+1} = Bw_n$, so $w_{n+1} = B^n w$, $w := u_1 - y$, $w \perp N$.

Let us prove $\lim_{n\to\infty} \|B^n w\| = 0$. If this is proved, then (6) follows, and Theorem 1 is proved. We have

$$I := \|B^n w\|^2 = \int_0^\infty \frac{a^{2n}}{(a + s)^{2n}} d(E_s w, w) = \int_{s\geq b} + \int_{0<s\leq b} := I_1 + I_2 \tag{10}$$

where $b > 0$ is a number and E_s is the resolution of the identity corresponding to the selfadjoint operator $T \geq 0$.

In the region $|s| \geq b$ one has

$$I_1 \leq q^n(b)\|w\|^2, \qquad 0 < q < 1, \qquad q = q(b) = \max_{|s|\geq b} \frac{a^2}{(a + s)^2}. \tag{11}$$

We estimate I_2 as follows:

$$I_2 \leq \int_{-b}^{b} d(E_s w, w) \tag{12}$$

uniformly with respect to n.

Since $w \perp N$, we have

$$\lim_{b\to 0} \int_0^b d(E_s w, w) = \|P_N w\|^2 = 0, \tag{13}$$

where $P_N = E_{+0} - E_0$ is the orthoprojector onto N, and

$$\lim_{b\to 0} \int_{-b}^{0} d(E_s w, w) = 0 \tag{14}$$

because $E_{s-0} = E_s$. Therefore, for an arbitrary small $\varepsilon > 0$, we choose $b > 0$ so small that $I_1 \leq \frac{\varepsilon}{2}$, and for fixed b we choose n so large that $q^n(b) \leq \frac{\varepsilon}{2}$. Then $I \leq \varepsilon$.

Theorem 1 is proved. $\square$

We could replace a by ia in the above arguments.

Proof of Theorem 3. Denote $f_\delta := g$ and $u_{a,\delta} := z := A^* Q_a^{-1} g$. The operator $A^* Q_a^{-1} = V Q^{\frac{1}{2}} Q_a^{-1}$, where V is a partial isometry, so $\|A^* Q_a^{-1}\| \leq \|Q^{\frac{1}{2}} Q_a^{-1}\| \leq \frac{1}{2\sqrt{a}}$. Thus, z is defined for any $g \in H$. We have

$$F(z+h) = F(z) + \|Ah\|^2 + a\|h\|^2 + 2Re[(Az - g, Ah) + a(z,h)] \qquad \forall h \in D(A). \tag{15}$$

If $z = A^* Q_a^{-1} g$, then

$$(Az - g, Ah) + a(z,h) = (QQ_a^{-1} g - g, Ah) + a(Q_a^{-1} g, Ah) = 0. \tag{16}$$

From (15) and (16) we obtain

$$F(z+h) \geq F(z) \qquad \forall h \in D(A) \tag{17}$$

and $F(z+h) = F(z)$ implies $h = 0$. Thus z is the unique global minimizer of F. Let us prove $A^* Q_a^{-1} = T_a^{-1} A^*$. Since both operators in this identity are bounded, it is sufficient to check that

$$A^* Q_a^{-1} \psi = T_a^{-1} A^* \psi \tag{18}$$

for all ψ in a dense subset of H. As such dense subset let us take $D(A^*)$. Denote $Q_a^{-1}\psi := g$. Then $\psi = Q_a g$. Equation (18) is equivalent to

$$T_a A^* g = A^* Q_a g, \text{ or } A^*AA^* g + aA^* g = A^*AA^* g + aA^* g,$$

which is an obvious identity. Reversing the steps, we obtain (18) for every $\psi \in D(A^*)$. Note that $\psi \in D(A^*)$ is equivalent to $g \in D(A^*AA^*)$, so that the above calculations are justified.

Theorem 3 is proved. $\square$

Equation (18) allows one to replace the term $T_a^{-1} A^* f$ in (5) by the term $A^* Q_a^{-1} f$ which is originally well defined for any $f \in H$.

3 Stable solution of (1).

Suppose that noisy data f_δ, $\|f_\delta - f\| \le \delta$, are given. We want to construct a stable approximation u_δ of y in the sense (4).

One way to do this is to use iterative process (5), with f_δ in place of f, and stop the iterations at the step $n = n(\delta)$, where $n(\delta)$ is properly chosen. Indeed, if we use the argument from the proof of Theorem 3, then we get $w_{n+1} = Bw_n + S(f_\delta - f)$. Thus $w_{n+1} = \sum_{j=0}^{n} B^j S(f_\delta - f) + B^n w$, where $w = u_1 - y$, $u_1 \perp N$, and

$$\|\sum_{j=0}^{n} B^j S(f_\delta - f)\| \le \frac{\delta(n+1)}{2\sqrt{a}} + \varepsilon(n) := \nu(\delta, n)$$

where $\lim_{n\to\infty} \varepsilon(n) = 0$, as we have demonstrated in the proof of Theorem 1.

It is clear that if $n = n(\delta)$ is chosen so that $\lim_{\delta\to 0} n(\delta) = \infty$ and $\lim_{\delta\to 0} \delta n(\delta) = 0$, then $\lim_{\delta\to 0} \nu(\delta, n(\delta)) = 0$. Therefore $u_\delta = u_{n(\delta)}$ satisfies (4).

4 Equation (1) with selfadjoint operator.

Assume that $A = A^*$, A is unbounded and A does not have a bounded inverse. Then one can use an analog of Theorem 1 in the form

$$u_{n+1} = Lu_n + g, \qquad L := ia(A + ia)^{-1}, \qquad g = (A + ia)^{-1} f, \tag{19}$$

where $a = const > 0$ and $u_1 \perp N := N(A)$ is arbitrary. Note that any element Ah, $\forall h \in D(A)$, is orthogonal to N since $A = A^*$. The minimal-norm solution y to (1) solves the equation $y = Ly + g$, so that the proof of Theorem 1 remains almost the same. So we get

Theorem 4. *If $A = A^*$ is unbounded, $a = const > 0$, and equation (1) is solvable, then (6) holds for the iterative process (19).*

Also an analog of the result of Section 3 holds.

5 DSM

In this Section we justify the dynamical systems method (DSM) for solving equation (1). The DSM theory is developed in [2], pp.41-70.

Theorem 5. *Assume that $f = Ay$, $y \perp N$, $N := N(A)$, A is a linear operator, closed and densely defined in H. Consider the problem*

$$\dot{u} = -u + T_{\varepsilon(t)}^{-1} A^* f, \qquad u(0) = u_0; \qquad \dot{u} := \frac{du}{dt}, \tag{20}$$

where $u_0 \in H$ is arbitrary, $T_\varepsilon = A^*A + \varepsilon I$, $\varepsilon = \varepsilon(t) > 0$ is a continuous function monotonically decaying to zero as $t \to \infty$, and $\int_0^\infty \varepsilon(s)ds = \infty$. Then problem (20) *has a unique solution $u(t)$ defined on $[0, \infty)$, there exists*

$$\lim_{t\to\infty} u(t) := u(\infty), \quad \text{and} \quad u(\infty) = y. \tag{21}$$

Proof. One has $T^{-1}_{\varepsilon(s)} A^* f = T^{-1}_{\varepsilon(s)} Ty$. Therefore

$$u(t) = u_0 e^{-t} + \int_0^t e^{-(t-s)} T^{-1}_{\varepsilon(s)} Ty\, ds. \tag{22}$$

The conclusion of Theorem 5 follows immediately from two lemmas:

Lemma 2. *If there exists $h(\infty) = \lim_{t\to\infty} h(t)$, then*

$$\lim_{t\to\infty} \int_0^t e^{-(t-s)} h(s)\, ds = h(\infty). \tag{23}$$

Lemma 3. *If $y \perp N := N(A)$, then*

$$\lim_{\varepsilon\to 0} T_\varepsilon^{-1} Ty = y. \tag{24}$$

The proof of Lemma 2 is simple and is left to the reader.

The proof of Lemma 3 is briefly sketched below::

$$T_\varepsilon^{-1} Ty - y = \int_0^\infty \left(\frac{s}{s+\varepsilon} - 1 \right) dE_s y = - \int_0^\infty \frac{\varepsilon}{s+\varepsilon} dE_s y.$$

Thus,

$$\lim_{\varepsilon\to 0} \|T_\varepsilon^{-1} Ty - y\|^2 = \lim_{\varepsilon\to 0} \int_0^\infty \frac{\varepsilon^2 d(E_s y, y)}{(s+\varepsilon)^2} = \|P_N y\|^2 = 0, \tag{25}$$

because $y \perp N$. The projector P_N is the orthogonal projector onto N.

Theorem 5 is proved. □

One can use Theorem 5, exactly as it is done in [2], for stable solution of equation (1) with noisy data: if f_δ is given in place of the exact data f, $\|f_\delta - f\| \le \delta$, then one solves problem (20) with f_δ in place of f, calculates its solution $u_\delta(t)$ at $t = t_\delta$, and proves that

$$\lim_{\delta\to 0} \|u_\delta(t_\delta) - y\| = 0 \tag{26}$$

if t_δ is suitably chosen. The stopping time t_δ can be uniquely determined, for example, by a discrepancy principle as shown in [2] for bounded operators A. The argument in [2] remains valid in the case of unbounded A without any changes.

References

[1] Kato, T., **Perturbation Theory for Linear Operators**, Springer Verlag, New York, 1984.

[2] Ramm, A. G., **Inverse Problems**, Springer Verlag, New York, 2005.